

Psycholinguistic Evidence for Severing Arguments from the Verb

SHINNOSUKE ISONO
University of Tokyo

YUKI HIROSE
University of Tokyo

1 Background: Mystery of Locality Effect

Many studies of human sentence processing assume that when a verb is input, preceding thematic arguments are retrieved from working memory (Gibson, 2000; Van Dyke and Lewis, 2003; Lewis et al., 2006; Vasishth et al., 2019; among many others). In (1), for example, the subject *neko-ga* ‘cat-NOM’ is assumed to be retrieved when the verb *tataita* ‘hit’ is input.

- (1) Neko-ga [inu-ga oikaketa] nezumi-o tataita.
cat-NOM [dog-NOM chased] mouse-ACC hit.
‘The cat hit the mouse that the dog chased.’

Two factors have been suggested in the literature to affect the difficulty of such retrieval: *similarity* and *locality*. Similarity refers to the feature-based similarity between the retrieval target and other elements in the sentence. According to the cue-based retrieval model (e.g., Van Dyke and Lewis, 2003), retrieval of the correct target becomes difficult in the presence of similar elements since the cues used in retrieval (such as [subject] and [animate]) are

overloaded. In (1), for example, *inu-ga* may interfere with *neko-ga* since both are subjects and animate. Locality, on the other hand, refers to the linear distance between the target and the retrieval site. Studies suggest that distant dependencies are more difficult to process because intervening elements interfere, or the representation of the target element decays over time (Gibson, 2000; Lewis and Vasishth, 2005¹; Van Dyke and McElree, 2011).

While similarity-based interference has received robust empirical support, evidence for locality effect is somewhat slippery (see Nakatani, 2021a, for a concise review). Many studies have observed reading slowdown in distant dependencies (Bartek et al., 2011; Levy et al., 2013; Ono and Nakatani, 2014; Safavi et al., 2016; among others), but some have failed to find one or even found an effect in the opposite direction (Konieczny, 2000; Vasishth and Lewis, 2006; Nakatani and Gibson, 2010).

Notably, unexpected results concentrate on simple argument-verb dependencies in verb-final languages. In Japanese, for example, the locality effect has been observed in dependencies between a *wh*-argument and a verb (plus complementizer); and between a negative polarity item and a negated verb (Ono and Nakatani, 2014; Nakatani, 2021a), but not in simple argument-verb dependencies without such additional complexity (Nakatani and Gibson, 2010). A similar contrast is observed in German (Levy and Keller, 2013). One explanation for this tendency is that orthogonal factors mask the underlying locality effect. Intervening elements may facilitate the processing of the verb by making it more predictable (Levy, 2008) or accessible (Vasishth and Lewis, 2006). Simple argument-verb dependencies in verb-final languages are particularly prone to these factors since the intervening elements are often arguments themselves and have strong ties to the verb. However, the lack of locality effect is not completely reducible to effects of expectation and accessibility, since the same tendency is observed even when these factors are controlled (Levy and Keller, 2013; Nakatani and Gibson, 2010).

For this reason, some studies suggest that head-directionality or the type of the dependency bears directly on the magnitude or even presence of locality effect (for head-directionality, see Levy and Keller, 2013; for dependency type, see Nakatani and Gibson, 2010; Ono and Nakatani, 2014; Nakatani, 2021a). This hypothesis is also consistent with the observation in English that verbs in relative clauses show a stronger locality effect than matrix verbs (Bartek et al., 2011). Based on these proposals, the current study investigates *how* certain structures evade locality effect by a reading experiment

¹ In the original ACT-R model by Lewis and Vasishth (2005), the major source of locality effect was retrieval of predicted heads. Since Lewis et al. (2006) and Vasishth and Lewis (2006), however, retrieval of arguments was featured in addition, and that seems to be the major concern of recent studies using the model (Vasishth et al., 2019).

in Japanese.

2 Experiment: Do We Really Retrieve Arguments at the Verb?³

2.1 Hypotheses and Design

One possible explanation for the lack of locality effect is that the parser adopts a strategy to avoid confusion due to retrieval of distant items. We test the following hypothesis (see Kimball, 1973 and Lewis, 1996 for similar proposals).

- (i) Arguments are cleared from working memory once their thematic dependency is complete.

This strategy would attenuate the locality effect if its major source is interference. For example, in (1), the embedded subject *inu-ga* could be cleared from memory once the embedded clause is complete; then, it would not interfere with the matrix subject *neko-ga* when the matrix verb is input. If interference is the major source of locality effect, this strategy would attenuate the effect.

Alternatively, we can cast doubt on the very idea that arguments must be retrieved at the verb (cf. Friedmann et al., 2008). It is possible that verb-final languages avoid immediate retrieval of arguments upon the input of the verb to evade heavy memory load due to retrieval of multiple arguments in the distance. The hypothesis can be stated as follows (here we limit the scope to Japanese since it is the only language we investigate).

- (ii) Retrieval of arguments does not take place at the verb in Japanese.

This is surprising if arguments are retrieved in order to form the dependency with the verb as an essential part of syntactic and/or semantic structure building. However, such dependencies may not be needed. According to the *constructivist* view in generative syntax (see Marantz, 2013), some or all thematic arguments are severed from the verb and instead introduced by functional heads, as shown in (2).⁴ Semantically, the verb provides a predicate P over the event e , and P is only indirectly related to the arguments x, y, z via e , as shown in (3).

(2) [VoiceP Agent [AppIP Goal [_{vP} Theme [_{√P} Root] v] AppI_{high}] Voice]

(3) $\lambda e. \text{Agent}(x, e) \& \text{Goal}(y, e) \& \text{Theme}(z, e) \& P(e)$

If such constructivist structure underlies sentence processing, there would be

³The experiment was originally presented in the 162nd Conference of Linguistic Society of Japan. The current study contains new statistical and theoretical analyses of the data.

⁴Exactly how arguments are realized is still debated within constructivism. To account for the current result, it suffices if dative arguments are severed from the verb.

Region →	1	2	3	4
A (+Interference, +Incompletion)	Taroo-ga Taroo-NOM	keezi-ni , detective-DAT	nigedasita [[ran.away]	doroboo-ga thief-NOM
B (+Interference, -Incompletion)	Taroo-ga, Taroo-NOM	keezi-ni [[detective-DAT	butukatta bumped]	doroboo-ga thief-NOM
C (-Interference)	Taroo-ga, Taroo-NOM	awatete [[in.a.hurry	nigedasita ran.away]	doroboo-ga thief-NOM

	5	6	7	8	9	10
A/B/C	zyunsa-ni policeman-DAT	ookina large	gin'irono silver	naihu-o knife-ACC	tukituketa held.against	moyooda-to] seems-that

	11	12	13	14	15	16
A/B	kanari very	hakkiri clearly	syoogensita-tame testified-so	soosa-ga investigation-NOM	kyuusokuni quickly	susunda. progressed
C	keezi-ni detective-DAT	hakkiri clearly	syoogensita-tame testified-so	soosa-ga investigation-NOM	kyuusokuni quickly	susunda. progressed

- 'The investigation progressed quickly as ...
- A Taroo testified to the detective very clearly that it seemed that the thief that ran away ...
- B Taroo testified very clearly that it seemed that the thief that bumped into the detective ...
- C Taroo testified to the detective clearly that it seemed that the thief that ran away in a hurry ...
held a large silver knife against the policeman.'

TABLE 1 Target sentences and their translations

no syntactic need to retrieve arguments at the verb. It would suffice to insert the lexical information of the verb to the Root node in the structure like (2), which is already built before the verb is revealed. Note that this is not to say that constructivism *prohibits* retrieval of arguments at the verb. It is possible that arguments are retrieved for other reasons (see Section 2.4). Hypothesis (ii) states, however, that that does not happen in Japanese.

These hypotheses are tested by a self-paced reading experiment using sentences like those shown in Table 1.⁵ The embedded verb at Region 9 (e.g., *tukituketa* 'held against') takes a dative argument at Region 5 (*zyunsa-ni* 'policeman-DAT'). In Conditions A and B, there is another dative NP (*keezi-ni* 'detective-DAT') in Region 2; in Condition C, the same NP is placed at Region 11, after the embedded verb. If arguments must be retrieved at the verb, as standardly assumed, there should be slowdown due to interference by the earlier dative NP in conditions A and B. If hypothesis (i) is correct, however, the earlier dative NP (*keezi-ni*) interferes only when its thematic dependency is not complete before the embedded verb is found, i.e., in Condition A. If hypothesis (ii) is correct, there would be no interference effect at all because there is no argument retrieval. Note that the argument structure of the embedded clause is the same across conditions, minimizing the effect of the aforementioned confounding factors such as expectation. The position and the lexical item of the critical region are also matched across conditions.

⁵ In Japanese, the left edge of an embedded clause is not explicitly marked. In the current target

2.2 Method

Participants. Forty-four native Japanese speakers from age 20 to 32 (mean: 21.4) participated. Each received 500 yen for participation. One participant was excluded from analysis because of a technical error.

Procedure. The experiment used the moving-window, self-paced reading paradigm (Just et al., 1982). It was conducted on the Ibex Farm website⁶, using the Ibex software by Alex Drummond. A session consisted of a written instruction, two practice trials, and seventy-two test trials. In each trial, a gaze point '+' was presented on the left, followed by a series of dashes that mask the words. When the participant pressed the space key, the next word appeared and the previous word (or the gaze point) was masked by a dash. After the sentence, a yes/no comprehension question followed. A short sentence (e.g., 'The thief held a silver knife against the detective.') was presented, and the participant pressed the F key if they thought the sentence was correct, or the J key if they thought it was not. The question sentence contained all the thematic arguments from either the matrix or the embedded clause to prevent participants from focusing on a particular element in advance, but if anything was incorrect, it was always the dative argument.

Materials. Twenty-four sets of target sentences were distributed into three lists in a Latin Square design. Each participant was assigned one of the lists. Forty-eight filler sentences were mixed with these target sentences in a pseudo-random order generated for each participant.

Analysis. Before any further analysis, two participants whose log-transformed reading time per region were 2.5 SDs below or above the mean (4.96 ln ms and 6.89 ln ms; mean: 6.06 ln ms; SD: 0.33) were excluded. Also, one participant whose comprehension accuracy was 2.5 SDs below the mean (55.6%; mean: 80.3%; SD: 9.7) was excluded, but this is the same participant as the one excluded on the basis of the reading time. The data from the remaining forty-one participants were submitted to further analysis. Furthermore, reading times 2.5 SDs below or above the mean, calculated by region and condition, were excluded (2.9% of all data points from target trials).

Comprehension accuracy and log-transformed reading times of the critical and spillover regions were analyzed by (generalized) linear mixed effects modelling (Baayen et al., 2008). Maximal models included fixed effects of [\pm Interference] (A/B vs. C) and [\pm Incompletion] (A vs. B). These effects were coded using Helmert contrasts (Schad et al., 2020), as shown in Table 2.

sentences, however, the reader should be able to detect a clause boundary when two nominative NPs are presented (Miyamoto, 2002). The verb in Region 3 and the dative NP in Region 5 should also help locate the boundary. A *tooten* (shown as comma in Table 1) is also added to the word before the boundary for the same purpose.

⁶<https://spellout.net/> The website was shut down after this experiment in September, 2020.

	Interference	Incompletion
A	1/3	1/2
B	1/3	-1/2
C	-2/3	0

TABLE 2 Coding scheme

	Correct %	SE
A	64.9%	2.9
B	67.4%	3.2
C	69.8%	3.3

TABLE 3 Comprehension accuracy by condition. SE: standard error.

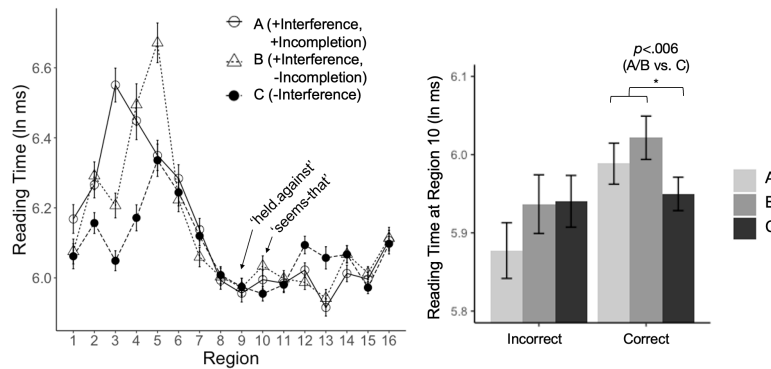


FIGURE 1 Left: Reading times by region and condition in trials with a correct answer to the comprehension question. Right: Reading times in Region 10 by condition and comprehension accuracy. Error bars indicate standard errors.

Comprehension accuracy and its interaction with the other factors were also included since the initial inspection of the data suggests different trends for trials with correct answer to the question (coded as 1/2) and those with incorrect ones (coded as -1/2). Random effects included intercepts and slopes for participants and items. Following Bates, Kliegl, et al. (2015), models were simplified by iteratively removing non-significant random effects.

The analysis was conducted using the R software (R Core Team, 2019). Models were fitted using the `glmer` and `lmer` functions in the `lme4` package (Bates, Mächler, et al., 2015), and p -values were estimated using the `lmerTest` package (Kuznetsova et al., 2017).

2.3 Result

Comprehension Accuracy. Table 3 shows comprehension accuracy by condition. No fixed effect reached significance.

Reading time. The left panel of Figure 1 shows reading times by region and condition in trials with a correct answer to the comprehension question. Data from trials with an incorrect answer showed similar trends, but there was one notable difference, described below.

	Estimate	SE	<i>t</i> -value	<i>p</i> -value	
Intercept	5.968	0.039	151.933	<.001	*
Interference	0.028	0.019	1.479	.140	
Incompletion	-0.040	0.032	-1.224	.233	
Accuracy	0.002	0.025	0.079	.937	
Interference:Accuracy	0.100	0.041	2.421	.016	*
Incompletion:Accuracy	0.031	0.049	0.636	.525	

TABLE 4 Model estimates for Region 10. SE: standard error. * indicates $p < .05$.

In Region 9 (embedded verb), no fixed effect reached significance. In Region 10 (spillover region), there was a significant interaction ($p < .02$) between [\pm Interference] and comprehension accuracy. Table 4 shows the estimated model for Region 10. Pairwise comparison revealed that the [+Interference] conditions were significantly slower than the [-Interference] condition only when the comprehension question was correctly answered ($p < .006$) (see the right panel of Figure 1). The trend was reversed (but not significant) for trials with an incorrect answer to the comprehension question.

2.4 Discussion

The result supports neither of the hypotheses we postulated initially. The effect in Region 10 indicates that the dative argument was retrieved at least in trials with a correct answer to the comprehension question, contrary to the prediction of hypothesis (ii). Furthermore, this interference effect was observed even in the [-Incompletion] condition (Condition B), which indicates that the interfering dative NP remained in working memory even after the dependency is complete. Thus hypothesis (i) was not supported either.

Rather, the interaction between [\pm Interference] and comprehension accuracy at Region 10 can be interpreted as follows. If the parser retrieves the dative argument at the verb, that results in *both* slowdown due to interference and better comprehension of the thematic relation. But the parser has another option, namely, not to retrieve the dative argument in order to avoid interference, at the cost of less accurate comprehension. In short, argument retrieval is optional.

This interpretation is consistent with constructivism. As discussed earlier, constructivism predicts that argument retrieval is not syntactically required at the verb. We then hypothesized that arguments are never retrieved at the verb in Japanese. But another possibility that stems from constructivism is that the parser commits to retrieval of arguments for non-syntactic reasons, e.g., for a better understanding of the event e (see (3)). The world knowledge associated with the verb (e.g., *tukituketa* ‘held against’) indicates the presence of a Goal argument, and the parser may look for one, hence the interference effect. But this is not necessary to form a grammatical representation of the sentence,

which is supposedly the minimal requirement in syntactic processing. Therefore the parser should be able to choose whether to retrieve arguments under the trade-off between better comprehension and cost of retrieval.

Importantly, this interpretation is also able to explain the lack of locality effect in previous studies. If argument retrieval is not due to syntactic requirement but for better comprehension, it can also take place in positions before the verb. For example, an argument may trigger retrieval of earlier arguments of the same clause, using the shared event variable as a cue.⁷ For this reason, the surface argument-verb distance does not determine the magnitude of locality effect at the verb as the argument may have been reactivated in an intermediate position. This is particularly the case in verb-final languages, where many arguments are introduced before the verb.

Under the constructivist view, locality effects that appear to be based on argument-verb distance should instead be explained by retrieval of predicted heads such as T, as Lewis and Vasishth (2005) did in the original ACT-R model (see footnote 1). This would explain locality effects in English subject-verb dependencies, and also locality effects in dependencies in verb-final languages that involve an additional predicted head (e.g., Neg in the case of negative concordance). In this regard, an important topic for future work is to investigate whether there are effects that can be attributed to the processing of argument-introducing heads assumed in the current hypothesis.

Before concluding, let us briefly discuss the data from the matrix verb. If locality affects simple argument-verb dependencies, Condition A should be slower than C at the matrix verb since the matrix dative argument (*keezi-ni*) is more distant. However, there was a significant effect in the opposite direction at Region 13 ($p < .001$). Since this effect continues from Region 12, this may be an effect of the scrambled word order (subject – sentential complement – indirect object) in Condition C, which becomes evident in Region 11. As this orthogonal effect is prevalent, the data is consistent with but not particularly supportive of the lack of locality effect in the dependency in question.

3 Conclusion

To summarize, the current experiment showed a significant interaction between retrieval interference of an argument and comprehension accuracy. This effect is consistent with the view that argument retrieval at the verb is optional. This view may also explain why simple argument-verb dependencies in verb-final structures often fail to show locality effect.

⁷Nakatani (2021b) recently suggested that dependents of the same clause are stored in a ‘dependency chain’, which is updated whenever a new dependent is added, and this is why thematic dependencies do not show locality effect. Although the structural assumption that Nakatani employs is different from ours, the underlying intuition seems similar.

Acknowledgments

We express our gratitude to the organizing committee, the anonymous reviewers, and all the visitors to our presentation at the conference.

References

- Baayen, R. H., Davidson, D. J., and Bates, D. M. 2008. Mixed-Effects Modeling with Crossed Random Effects for Subjects and Items. *Journal of Memory and Language* 59(4):390–412.
- Bartek, B., Lewis, R. L., Vasishth, S., and Smith, M. R. 2011. In Search of On-Line Locality Effects in Sentence Comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37(5):1178–1198.
- Bates, D., Kliegl, R., Vasishth, S., and Baayen, R. H. 2015. Parsimonious Mixed Models. arxiv:1506.04967 [stat.me].
- Bates, D., Mächler, M., Bolker, B., and Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1):1–48.
- Friedmann, N., Taranto, G., Shapiro, L. P., and Swinney, D. 2008. The Leaf Fell (the Leaf): The Online Processing of Unaccusatives. *Linguistic Inquiry* 39(3):355–377.
- Gibson, E. 2000. The Dependency Locality Theory: A Distance-Based Theory of Linguistic Complexity. *Image, Language, Brain*. Cambridge, eds. A. P. Marantz, Y. Miyashita, and W. O’Neil, 95–126. MA: MIT Press.
- Just, M. A., Carpenter, P. A., and Woolley, J. D. 1982. Paradigms and Processes in Reading Comprehension. *Journal of Experimental Psychology: General* 111(2):228–238.
- Kimball, J. 1973. Seven Principles of Surface Structure Parsing in Natural Language. *Cognition* 2(1):15–47.
- Konieczny, L. 2000. Locality and Parsing Complexity. *Journal of Psycholinguistic Research* 29(6):627–45.
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13):1–26.
- Levy, R. 2008. Expectation-Based Syntactic Comprehension. *Cognition* 106(3):1126–1177.
- Levy, R., Fedorenko, E., and Gibson, E. 2013. The Syntactic Complexity of Russian Relative Clauses. *Journal of Memory and Language* 69(4):461–495.
- Levy, R. and Keller, F. 2013. Expectation and Locality Effects in German Verb-Final Structures. *Journal of Memory and Language* 68(2):199–222.
- Lewis, R. L. 1996. Interference in Short-Term Memory: The Magical Number Two (or Three) in Sentence Processing. *Journal of Psycholinguistic Research* 25:93–115.
- Lewis, R. L. and Vasishth, S. 2005. An Activation-Based Model of Sentence Processing as Skilled Memory Retrieval. *Cognitive Science* 29(3):375–419.
- Lewis, R. L., Vasishth, S., and Van Dyke, J. A. 2006. Computational Principles of Working Memory in Sentence Comprehension. *Trends in Cognitive Science* 10(10):447–454.

- Marantz, A. 2013. Verbal Argument Structure: Event and Participants. *Lingua* 130:152–168.
- Miyamoto, E. T. 2002. Case Markers as Clause Boundary Inducers in Japanese. *Journal of Psycholinguistic Research* 31(4):307–347.
- Nakatani, K. 2021a. Locality Effects in the Processing of Negative-Sensitive Adverbials in Japanese. *The Joy and Enjoyment of Linguistic Research: A Festschrift for Takene Ito*, eds. R. Okabe, J. Yashima, Y. Kubota, and T. Isono. 462–472. Tokyo: Kaitakusya.
- Nakatani, K. 2021b. Integrating Different Types of Long-Distance Dependencies: A Case Study of a Negative Polarity Dependency in Japanese. Talk at *International Symposium on Issues in Japanese Psycholinguistics from Comparative Perspective*.
- Nakatani, K. and Gibson, E. 2010. An On-Line Study of Japanese Nesting Complexity. *Cognitive Science* 34(1):94–112.
- Ono, H. and Nakatani, K. 2014. Integration Costs in the Processing of Japanese *Wh*-Interrogative Sentences. *Studies in Language Sciences* 13:13–31.
- R Core Team. 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Safavi, M. S., Husain, S., and Vasishth, S. 2016. Dependency Resolution Difficulty Increases with Distance in Persian Separable Complex Predicates: Evidence for Expectation and Memory-Based Accounts. *Frontiers in Psychology* 7(403).
- Schad, D. J., Vasishth, S., Hohenstein, S., and Kliegl, R. 2020. How to Capitalize on A Priori Contrasts in Linear (Mixed) Models: A Tutorial. *Journal of Memory and Language* 110:104038.
- Van Dyke, J. A. and Lewis, R. L. 2003. Distinguishing Effects of Structure and Decay on Attachment and Repair: A Cue-Based Parsing Account of Recovery from Misanalyzed Ambiguities. *Journal of Memory and Language* 49(3):285–316.
- Van Dyke, J. A. and McElree, B. 2011. Cue-Dependent Interference in Comprehension. *Journal of Memory and Language* 65(3):247–263.
- Vasishth, S. and Lewis, R. L. 2006. Argument-Head Distance and Processing Complexity: Explaining Both Locality and Antilocality Effects. *Language* 82(4):767–794.
- Vasishth, S., Nicenboim, B., Engelmann, F., and Burchert, F. 2019. Computational Models of Retrieval Processes in Sentence Processing. *Trends in Cognitive Sciences* 23(11):968–982.